

Predição de Peso da Tilápia do Nilo por Meio do Fenótipo Usando Redes Neurais Convolucionais

Henrique Yoshiharu Kajihara
Universidade Estadual de Maringá (UEM)
henriqueykajihara@gmail.com

Josiane Melchiori Pinheiro
Universidade Estadual de Maringá (UEM)
jmpferreira@uem.br

Abstract

Biometric analysis in fish farms, such as weighing and measuring Nile tilapia, is a manual and stressful procedure for the animals and incurs significant costs, leading to growth delays and stock losses. This study investigates the application of Convolutional Neural Networks (CNNs) to estimate tilapia weight from images of its phenotype, with the aim of automating and optimizing this process. Using an existing image dataset, several CNN models were trained and evaluated. The results are promising, with some configurations achieving discrepancies of only 45–50 grams between predicted and actual weight, and a coefficient of determination (R^2) of 0.97. It is concluded that it is feasible to develop an efficient and accurate system to predict the weight of Nile tilapia using CNNs, with potential to increase productivity and reduce costs in fish farming.

Palavras-chave

Inteligência Artificial (IA), Redes Neurais Convolucionais (CNNs), Visão Computacional, Tilápia do Nilo, Piscicultura.

1 Introdução

A utilização da tecnologia no agronegócio tem se tornado cada vez mais frequente, impulsionada pelo avanço de ferramentas inteligentes, sensores, georreferenciamento e métodos de análise automatizada. Esses recursos, frequentemente associados ao conceito de agricultura 4.0, têm proporcionado maior eficiência e precisão às etapas produtivas. No contexto da piscicultura, esse movimento tecnológico começa a ganhar força, especialmente no cultivo da Tilápia do Nilo, espécie de grande importância econômica para o Brasil.

Segundo dados da Associação Brasileira de Piscicultura (PeixeBR), o Brasil ocupa atualmente a quarta posição no ranking mundial de produção de tilápia, sendo responsável por aproximadamente 65% da produção nacional de peixes cultivados. Entre 2014 e 2023, a produção aquícola nacional aumentou de 578.800 para 887.029 toneladas, representando um crescimento superior a 50% no período [1]. Essa expansão evidencia a relevância da cadeia produtiva da tilápia e reforça a necessidade de aprimorar técnicas de manejo, monitoramento e tomada de decisão.

No manejo convencional, a biometria é uma etapa essencial para o acompanhamento do desenvolvimento dos peixes. Entretanto, tal processo ainda é majoritariamente manual, envolvendo a captura seletiva dos animais com tarrafa, manipulação direta, e muitas vezes o uso de sedativos para evitar que os peixes se debatam ou sofram lesões durante a pesagem [2, 3]. Esse procedimento exige pelo menos três operadores e consome tempo, aumenta o risco de contaminação e gera estresse significativo nos peixes.

O estresse fisiológico decorrente desse manejo é amplamente documentado na literatura zootécnica. Em [2] os autores descrevem

três níveis de resposta: os efeitos primários, associados à liberação de hormônios como catecolaminas e corticosteróides; os efeitos secundários, que incluem mudanças metabólicas como aumento de glicose, redução de glicogênio e acúmulo de lactato; e os efeitos terciários, que afetam diretamente o desempenho produtivo, reduzindo o crescimento, enfraquecendo o sistema imunológico e aumentando a mortalidade [2, 3]. Para sistemas de criação intensiva, esses impactos comprometem a sustentabilidade econômica e o bem-estar animal.

Nesse cenário, métodos automatizados e não invasivos de estimativa de peso tornam-se alternativas altamente promissoras. Abordagens baseadas em visão computacional permitem reduzir drasticamente a manipulação física dos peixes, minimizando o estresse e tornando o processo mais rápido e eficiente. Modelos de aprendizado de máquina e, em especial, Redes Neurais Convolucionais (CNNs - *Convolutional Neural Network*) têm se destacado pela capacidade de extrair padrões relevantes diretamente das imagens, alcançando performances superiores a métodos tradicionais baseados em regressões manuais [4, 5].

Além da melhora produtiva, a adoção dessas tecnologias favorece práticas mais éticas e sustentáveis, atendendo às exigências de consumidores e mercados que valorizam responsabilidade socioambiental [6]. A modernização da piscicultura por meio da automação e da inteligência artificial permite aprimorar o uso de insumos, melhorar a precisão do manejo e reduzir a mortalidade associada ao estresse da manipulação convencional.

Diante desse cenário, este trabalho investiga o uso de CNNs para a predição do peso da Tilápia do Nilo por meio de imagens do fenótipo. Utilizando um banco de dados real coletado em ambiente experimental, busca-se avaliar a viabilidade de um sistema automatizado capaz de substituir parte do manejo biométrico tradicional, contribuindo para a evolução tecnológica da piscicultura em direção a processos mais eficientes, seguros e sustentáveis.

2 Trabalhos Relacionados

A aplicação de visão computacional para estimar características fenotípicas não é uma abordagem nova, mas sua aplicação em piscicultura é menos explorada.

Ma et al. [7] revisaram mais de 70 pesquisas sobre o uso de aprendizado de máquina (incluindo regressão linear e redes neurais) em imagens 2D e 3D para estimar o peso de suínos e bovinos. Eles destacam que a abordagem reduz o estresse animal e permite a coleta massiva de dados, alcançando erros percentuais tão baixos quanto 4,87% em alguns casos. No entanto, apontam desafios como a necessidade dos animais permanecerem parados para imagens 3D e o alto custo de sensores de boa qualidade.

Lee, Lee e Cho [8] utilizaram redes neurais profundas em imagens 2D para estimar o peso de bovinos. O estudo comparou métodos de segmentação, utilizando Mask R-CNN (segmentação totalmente

supervisionada) e alcançando uma margem de erro de 5,06% em um conjunto de 43 animais.

Inspirado por esses avanços, este trabalho aplica técnicas similares de redes neurais profundas para previsão do peso das tilápias a partir de imagens, adaptando as estratégias para os desafios específicos da piscicultura.

Apesar dos resultados promissores em suínos e bovinos, observa-se uma lacuna quando se trata de espécies aquáticas. A maior parte dos trabalhos ainda se concentra em ambientes controlados, com câmeras e sensores dedicados, o que dificulta a transferência direta dessas soluções para a realidade de sistemas aquícolas de baixo custo. Além disso, contribuições como a de Cardoso et al. [9] exploram análise de imagens para estimativa de rendimento de carcaça, mas não abordam explicitamente a predição de peso a partir de CNNs. Assim, o presente estudo contribui ao explorar um cenário intermediário, em que se utilizam imagens coletadas em ambiente experimental realista, com iluminação e posicionamento não totalmente padronizados, aproximando-se das condições encontradas por produtores.

3 Metodologia

O desenvolvimento foi conduzido em etapas sistemáticas, desde a aquisição dos dados até a avaliação do modelo.

3.1 Materiais e Ambiente

Banco de Imagens: foram utilizadas 5975 imagens no formato JPG de 1195 tilápias (5 imagens de cada animal em ângulos diversos). Os dados foram obtidos na base experimental da PEIXEGEN (Núcleo de Pesquisa, manejo, melhoramento e genética molecular em piscicultura de água doce, pertencente ao Departamento de Zootecnia do Centro de Ciências Agrárias da UEM), provenientes do estudo de Cardoso et al. [9]. Os procedimentos experimentais do estudo de origem foram aprovados pela Comissão de Ética no Uso de Animais da Universidade Estadual de Maringá (CEUA/UEM), sob o protocolo CEUA nº 9452160720. As imagens originais foram capturadas com fundo azul uniforme e uma régua de 30 cm imediatamente abaixo dos peixes, permitindo a utilização da escala como referência de tamanho. O estudo de origem empregou imagens RGB e em profundidade para estimar o rendimento de filé [9]. Neste trabalho foram utilizadas apenas as imagens RGB.

A Figura 1 ilustra o processo de coleta das tilápias para registro das imagens, em ambiente de campo, enquanto a Figura 2 apresenta um exemplo de imagem bruta utilizada como entrada da CNN, evidenciando o padrão de fundo e a referência de escala adotados.

3.2 Divisão dos dados e protocolo de avaliação

Para evitar *data leakage* e inflar artificialmente as métricas, a divisão dos dados foi feita no **nível do indivíduo**: todas as 5 imagens de um mesmo peixe permaneceram sempre no mesmo subconjunto (treino, validação ou teste). Foram utilizados **1.000 indivíduos** (totalizando 5.000 imagens) para treinamento e validação, e **195 indivíduos** (975 imagens) foram reservados para um **teste independente**.

Uma fração do conjunto de treinamento foi separada como **validação** fixa para monitorar a convergência e apoiar escolhas de configuração (como o número de épocas e o comportamento dos



Figura 1: Processo de captura das imagens no ambiente experimental, com uso de tarrafa para coleta seletiva das tilápias.



Figura 2: Exemplo de imagem bruta de Tilápia do Nilo utilizada como entrada da CNN.

otimizadores), enquanto o conjunto de teste permaneceu completamente isolado até a avaliação final.

Devido à alta complexidade computacional inerente ao modelo proposto e ao tempo demandado para a execução dos experimentos, a validação cruzada não foi implementada neste estudo. Embora a validação cruzada seja reconhecida como uma técnica estatística robusta para estimar a capacidade de generalização, ela requer múltiplas execuções completas do processo de treinamento e validação, o que implicaria um custo computacional proibitivo no contexto deste trabalho. Diante disso, optou-se por uma abordagem mais viável, utilizando um conjunto de validação fixo e um conjunto de teste independente para avaliar o desempenho do modelo. Embora essa metodologia permita uma análise inicial consistente, a implementação da validação cruzada em estudos futuros pode proporcionar uma avaliação mais abrangente e confiável, mitigando possíveis vieses associados à divisão dos dados e aprimorando a robustez das conclusões obtidas.

Ferramentas: o modelo foi implementado em Python 3.12, utilizando TensorFlow 2.18.0, OpenCV, Pandas e Matplotlib. Essas bibliotecas foram escolhidas por oferecerem suporte consolidado à

construção de redes neurais profundas, processamento de imagens e visualização de resultados.

Ambiente experimental: os experimentos foram executados em um computador com processador Intel Core i7-12650H e 32 GB de RAM DDR4. Não foi utilizada GPU dedicada, o que motivou a adoção de uma arquitetura de CNN de complexidade moderada, visando equilibrar desempenho e custo computacional.

Análise exploratória das imagens: antes do treinamento, foi realizada uma inspeção visual das imagens e dos metadados associados, com o objetivo de identificar padrões de iluminação, variações de enquadramento e presença de reflexos. Observou-se que, embora o cenário de captura fosse controlado (fundo azul e régua fixa), havia variações relevantes na posição dos peixes em relação à câmera e na incidência de luz, fatores que impactaram diretamente a qualidade do pré-processamento e a estabilidade do treinamento.

Preparação dos rótulos: os pesos das tilápias variavam em uma faixa aproximada entre 400g e 2000g. Para facilitar o treinamento e reduzir a sensibilidade a diferenças de escala, os valores foram normalizados por meio de uma transformação do tipo *min-max*. Essa normalização contribuiu para gradientes mais estáveis nas primeiras épocas, facilitando a convergência dos otimizadores Adam e Nadam.

Divisão dos conjuntos de dados: a separação entre os conjuntos de treino e teste foi realizada de forma aleatória, assegurando a independência entre as amostras utilizadas para aprendizagem e avaliação do modelo. Para evitar vazamento de informação, cada indivíduo (peixe) teve todas as suas imagens alocadas integralmente a apenas uma das partições. O conjunto de treinamento foi composto por 1.000 peixes (5.000 imagens) e o conjunto de teste por 195 peixes (975 imagens), preservando a diversidade de pesos em ambas as partições.

3.3 Forma de Aprendizagem Utilizada

O presente trabalho utiliza um modelo baseado em **aprendizado supervisionado**, no qual cada imagem de entrada possui um rótulo associado, correspondente ao peso real da tilápia registrado no momento da biometria. De acordo com Mitchell (1997) e Russell e Norvig (2016), o aprendizado supervisionado consiste em treinar um modelo para aprender uma função que mapeia entradas para saídas com base em exemplos rotulados, sendo amplamente utilizado em tarefas de regressão como a proposta neste estudo.

O fluxo completo do modelo segue exatamente o padrão descrito na monografia: uma etapa de **entrada** com imagens no formato $128 \times 128 \times 3$ (RGB); um **processamento** intermediário composto por camadas convolucionais e de pooling que extraem características visuais importantes; e uma **saída** constituída por um único valor contínuo representando a estimativa de peso da tilápia.

A escolha pelo aprendizado supervisionado se justifica porque o conjunto de dados possui medidas precisas de peso, permitindo o uso de métricas como MAE (*Mean Absolute Error*) e R^2 (Coeficiente de Determinação), que facilitam a avaliação do desempenho do modelo. Esse tipo de abordagem também é consistente com estudos semelhantes, como os de Zhang et al. (2021)[10], que utilizaram regressão supervisionada para estimar peso em suínos a partir de imagens.

3.4 Arquitetura do Modelo Proposto

Foi desenvolvida uma CNN para regressão, projetada para processar imagens de entrada de $128 \times 128 \times 3$ pixels (RGB). A arquitetura, resumida na Tabela 1, foi baseada em estudos de visão computacional para tarefas similares.

Tabela 1: Resumo da arquitetura da CNN proposta.

Camada	Configuração	Função de Ativação
Entrada	Imagem $128 \times 128 \times 3$	
Convolutacional 1	32 filtros 3×3	ReLU
MaxPooling 1	2×2	-
Convolutacional 2	64 filtros 3×3	ReLU
MaxPooling 2	2×2	-
Convolutacional 3	128 filtros 3×3	ReLU
MaxPooling 3	2×2	-
Convolutacional 4	256 filtros 3×3	ReLU
MaxPooling 4	2×2	-
Flatten	-	-
Dense 1	512 neurônios	ReLU
Dropout	Taxa 0.5	-
Dense 2	256 neurônios	ReLU
Dropout	Taxa 0.5	-
Dense 3	128 neurônios	ReLU
Saída (Regressão)	1 neurônio	Linear

As camadas convolucionais extraem características visuais, enquanto as camadas de *pooling* reduzem a dimensionalidade espacial. A Regularização L2 foi aplicada às camadas densas e o *Dropout* foi usado para mitigar *overfitting*. A opção por uma arquitetura relativamente compacta também se justifica pela ausência de GPU no ambiente experimental, tornando inviável, neste momento, o uso de modelos significativamente mais profundos, como ResNets ou EfficientNets [11].

3.5 Detalhes da Rede Neural Convolucional

A CNN desenvolvida foi projetada especificamente para processar dados estruturados em grade, como imagens RGB de tilápias, seguindo os fundamentos clássicos apresentados por Goodfellow et al. (2016) e Mitchell (1997). As CNNs são compostas por camadas interconectadas capazes de extrair características hierárquicas — desde padrões simples, como bordas, até estruturas mais complexas, como formas e regiões anatômicas do peixe.

A arquitetura adotada é composta por quatro camadas convolucionais, cada uma seguida de uma camada *MaxPooling*. Os filtros utilizados tiveram tamanhos de 3×3 , com quantidades crescentes (32, 64, 128 e 256), conforme recomendado pela literatura para captar gradualmente padrões mais complexos nas imagens.

Após a etapa convolutacional, a camada *Flatten* transforma o mapa de características em um vetor unidimensional que alimenta um bloco denso composto por três camadas totalmente conectadas (512, 256 e 128 neurônios) com ativação *ReLU* e regularização via *Dropout*.

A camada de saída possui ativação linear, apropriada para tarefas de regressão.

A função de ativação *ReLU* foi utilizada para introduzir não linearidade e permitir que a rede capture relações complexas entre as características extraídas das imagens. Já o *MaxPooling* otimiza o processamento ao reduzir as dimensões espaciais e manter apenas as características principais.

3.6 Treinamento e Estratégia de Otimização

O processo de treino adotou o **Erro Absoluto Médio (MAE)** como função de perda, por ser uma métrica amplamente utilizada em problemas de regressão e por permitir interpretação direta em unidades físicas do problema. Além das métricas finais, todas as configurações foram executadas com diferentes **sementes de inicialização** para avaliar robustez e reduzir a dependência de uma única execução.

Foram avaliados dois otimizadores: **Adam** e **Nadam**. O Adam é conhecido pela boa estabilidade e capacidade de lidar com gradientes ruidosos, enquanto o Nadam incorpora o *Nesterov Accelerated Gradient*, o que proporciona maior suavidade na convergência e melhor desempenho em casos específicos, o que pode proporcionar maior suavidade na convergência e ganhos de desempenho em determinadas configurações.

Além disso, foi adotada uma estratégia de redução dinâmica da taxa de aprendizado após 50 épocas, aplicando um fator de 0,1 para permitir refinamentos finais do modelo. Esse mecanismo de controle da taxa de aprendizado ajudou a estabilizar a curva de treinamento e evitar oscilações nas últimas épocas.

Para mitigar o tamanho relativamente limitado do banco de imagens, foi utilizado aumento de dados (*data augmentation*) incluindo rotações, zoom, translações e espelhamentos horizontais. Essa estratégia ampliou significativamente a variedade de exemplos vistos pela rede, contribuindo para maior capacidade de generalização.

Por fim, o fluxo geral do modelo segue três etapas fundamentais: (1) entrada das imagens originais em formato .jpg, apenas ajustadas ao formato de entrada da rede neural; (2) extração de características via convoluções e *pooling*; e (3) regressão do peso pelas camadas densas.

4 Resultados obtidos

4.1 Pré-processamento

Embora as imagens tenham sido obtidas em um cenário parcialmente controlado (fundo azul e régua de referência), foi avaliado um *pipeline* de pré-processamento como etapa complementar, com o objetivo de padronizar as entradas e facilitar a extração de características visuais relevantes. Para isso, foram testadas operações com OpenCV, incluindo ajuste de brilho/contraste, conversão para o espaço HSV (*Hue, Saturation, Value* - Matiz, Saturação e Valor) e aplicação de máscaras para tentativa de segmentação do animal. No entanto, a abordagem mostrou-se instável e, em diversos casos, comprometeu a integridade da região de interesse, motivo pelo qual os resultados reportados neste artigo priorizam os modelos treinados diretamente com as imagens originais.

Como ilustrado na Figura 3, o *pipeline* de pré-processamento mostrou-se falho para lidar com a variabilidade presente nas imagens, falhando em segmentar corretamente o peixe em um número significativo de casos. Em diversas situações, partes essenciais do

animal foram removidas juntamente com o fundo, enquanto, em outras, o fundo não foi devidamente eliminado. Esses problemas foram atribuídos principalmente às variações de iluminação, aos reflexos na superfície corporal do peixe e às diferenças de posicionamento em relação à câmera. Como consequência direta dessas falhas, os modelos treinados com imagens pré-processadas apresentaram desempenhos significativamente inferiores, indicando que a CNN treinada diretamente com as imagens originais foi mais eficaz em lidar com a complexidade e a variabilidade dos dados.

A Tabela 2 apresenta os melhores desempenhos obtidos com a aplicação de técnicas de pré-processamento das imagens. Observa-se que, mesmo nas configurações mais favoráveis, os valores de erro médio absoluto (MAE) permaneceram elevados e o coeficiente de determinação (R^2) foi inferior ao observado nos modelos treinados diretamente com as imagens originais, indicando que o pré-processamento não contribuiu para a melhoria do desempenho e limitou a capacidade de generalização do modelo.

Tabela 2: Melhores resultados obtidos com uso de pré-processamento das imagens.

Otimizador	Épocas	Semente	MAE (g)	R^2
Adam	40	30	61,56	0,94
Nadam	20	100	256,46	0,12



Figura 3: Comparação da imagem original (esquerda) e o resultado do pré-processamento com falha (direita), que removeu partes do peixe.

4.2 Resultados dos Modelos

Os melhores resultados foram consistentemente obtidos sem o pré-processamento. A Tabela 3 sumariza os melhores desempenhos encontrados para os otimizadores Adam e Nadam.

Tabela 3: Melhores resultados obtidos (sem pré-processamento).

Otimizador	Épocas	Semente	MAE (g)	R^2
Adam	30	100	51,28	0,96
Nadam	30	30	45,00	0,97
Nadam	40	50	46,64	0,96
Nadam	50	20	48,12	0,96

4.2.1 O que muda entre Adam e Nadam. A comparação entre os resultados apresentados nas Tabelas 2 e 3 foi conduzida alterando exclusivamente o **otimizador utilizado durante o treinamento da rede neural**. Todos os demais elementos do experimento foram mantidos constantes, incluindo a arquitetura da CNN, o conjunto de dados utilizado, o esquema de divisão por indivíduo, o uso do mesmo conjunto de teste independente, bem como as configurações de treinamento, como número de épocas e sementes de inicialização.

No experimento correspondente à Tabela 2 foi utilizado o otimizador **Adam**, amplamente empregado em tarefas de aprendizado profundo devido à sua capacidade de adaptação da taxa de aprendizado e estabilidade durante o processo de treinamento. Já no experimento da Tabela 3 foi utilizado o otimizador **Nadam**, que consiste em uma variação do Adam que incorpora o momento de Nesterov, podendo proporcionar uma convergência mais estável e ajustes mais refinados dos pesos da rede.

Dessa forma, as diferenças observadas nas métricas de desempenho, como o MAE e o coeficiente de determinação (R^2), podem ser atribuídas principalmente ao comportamento distinto desses algoritmos de otimização durante o processo de aprendizado da rede neural, uma vez que toda a estrutura experimental permaneceu inalterada.

4.2.2 Interpretação prática do erro (MAE). No melhor cenário, o modelo obteve MAE de aproximadamente 45 g. Considerando o conjunto de teste, o peso médio por imagem é de aproximadamente 1061 g, o que implica um erro relativo médio da ordem de 4–5%. Esse nível de erro tende a ser **aceitável para aplicações de manejo** [7], cujo objetivo principal é acompanhar tendências de crescimento e ajustar o arraçamento¹ por lote, especialmente quando comparado ao custo operacional e ao estresse associados à pesagem manual.

Também é importante destacar que a distribuição de pesos observada no experimento (vide faixa aproximada de 400–2000 g no gráfico da Figura 4) cobre uma etapa relevante da engorda, tornando a avaliação representativa para cenários práticos.

A Figura 4 ilustra a forte aderência entre os pesos reais e os pesos preditos pelo melhor modelo no conjunto de teste. Embora existam algumas discrepâncias, a maioria das predições segue de perto os valores reais.

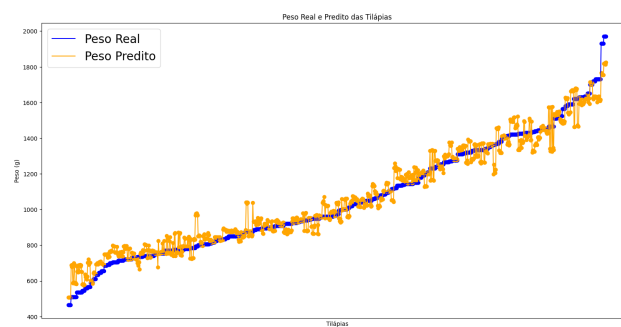


Figura 4: Comparativo dos pesos reais (azul) e preditos (laranja) usando Nadam (30 épocas, 30 sementes, sem pré-processamento).

4.3 Análise detalhada dos erros

Uma análise complementar da distribuição dos erros absolutos mostrou que a maior parte das predições se concentra em uma faixa inferior a 60 g de diferença em relação ao peso real, com poucos *outliers* acima de 100 g. Ao inspecionar qualitativamente tais casos, verificou-se que boa parte deles estava associada a imagens com reflexos muito intensos na pele do peixe, inclinações acentuadas ou enquadramento parcial do corpo, o que sugere que a CNN depende fortemente da integridade da silhueta para realizar estimativas mais precisas.

Considerando o conjunto de teste como um todo, a soma dos pesos reais foi de aproximadamente 1034,33 kg, enquanto a soma dos pesos preditos foi de 1045,92 kg, resultando em diferença global de cerca de -11,60 kg entre ambos. O desvio padrão dos erros absolutos foi da ordem de 38,09 g e o desvio padrão dos erros percentuais de 4,82%. Esses valores indicam que, embora exista uma leve tendência à superestimação dos pesos, a maioria das predições permanece em uma faixa considerada aceitável para aplicações de campo, especialmente quando comparada às variações inerentes ao processo de pesagem manual.

Além disso, observou-se uma leve tendência de superestimação dos pesos em animais mais leves e subestimação em animais mais pesados, indicando uma possível regressão à média. Esse comportamento é comum em problemas de regressão e pode ser mitigado em trabalhos futuros com arquiteturas mais sofisticadas ou com a inclusão de informações biométricas adicionais, como o comprimento do peixe.

4.4 Curvas de aprendizado e estabilidade dos otimizadores

As curvas de aprendizado revelaram convergência relativamente rápida nas primeiras épocas, seguida por um *plateau* nas etapas finais. Em algumas configurações sem *Dropout*, sinais de *overfitting* começaram a aparecer com o aumento do número de épocas, reforçando a importância do uso combinado de regularização L2 e *Dropout* nas camadas densas.

Comparando os otimizadores Adam e Nadam, este último apresentou trajetórias de perda mais suaves e estabilidade superior em cenários com maior variabilidade luminosa, o que está em linha

¹Arraçãoamento de peixes é a alimentação controlada e estratégica na piscicultura, visando o crescimento uniforme e a saúde dos peixes, fornecendo a ração correta, na quantidade certa, no momento certo e distribuída por todo o tanque para evitar competição e garantir que todos comam, ajustando a frequência e porção conforme idade, espécie e temperatura da água para otimizar a conversão alimentar e o desenvolvimento.

com a literatura sobre o uso de gradiente acelerado de Nesterov em superfícies de perda não convexas [12]. Em particular, o Nadam mostrou-se menos sensível a escolhas específicas de semente de inicialização, produzindo resultados mais consistentes entre diferentes execuções.

4.5 Discussão

O desempenho superior do Nadam sobre o Adam pode ser atribuído à sua combinação do Adam com o *Nesterov Accelerated Gradient* (NAG) [12]. O Nadam antecipa a direção do gradiente, o que proporciona maior estabilidade em superfícies de perda ruidosas, como as geradas pela variabilidade de iluminação e posicionamento das tilápias nos dados de treino.

Intervalos de confiança e comparação estatística: como o experimento foi repetido com múltiplas sementes, é possível reportar, além do melhor caso, medidas de variabilidade (por exemplo, média e desvio padrão do MAE), permitindo avaliar a estabilidade do modelo e a consistência dos resultados entre diferentes execuções. Comparações entre configurações podem ser realizadas considerando essa variabilidade, reduzindo o risco de conclusões baseadas apenas em diferenças pontuais.

Baselines e estratificação por faixa de peso: para contextualizar o ganho obtido pela CNN, trabalhos futuros podem incluir baselines simples (como regressão linear a partir de medidas extraídas), permitindo comparar modelos lineares e não lineares em tarefas de regressão visual [11]. Além disso, análises estratificadas por faixa de peso podem ser realizadas para verificar se o erro relativo se mantém estável em animais mais leves e mais pesados.

Comparação com o manejo tradicional: quando comparado ao processo manual de pesagem por amostragem, que exige, no mínimo, dois operadores, equipamentos específicos e, em alguns casos, sedação, o modelo proposto oferece um ganho potencial em termos de segurança, tempo e bem-estar animal. Ainda que a solução atual exija a captura do peixe para obtenção da imagem, abre-se caminho para futuramente integrar o modelo a sistemas de captura automática de imagens, inclusive subaquáticas, reduzindo ainda mais a intervenção humana.

Impacto na cadeia produtiva: em um cenário de produção intensiva, estimativas de peso mais frequentes permitem ajustes finos na taxa de arraçoamento e na densidade de estocagem, com impacto direto em custos de ração e sanidade dos peixes. Um sistema baseado neste modelo poderia, por exemplo, ser acoplado a uma esteira de classificação ou a um módulo de triagem automática, fornecendo estimativas em tempo quase real para apoio à tomada de decisão.

Limitações observadas: entre as principais limitações do trabalho destacam-se: (i) a ausência de variáveis físicas adicionais (como comprimento corporal) como entradas do modelo, (ii) a heterogeneidade das condições de iluminação e enquadramento das imagens, que prejudicou o pré-processamento, e (iii) o tamanho relativamente limitado da base em termos de diversidade de cenários de captura. Esses fatores indicam que, embora os resultados sejam encorajadores, ainda é necessário validar a abordagem em bancos de dados maiores, mais variados e coletados em diferentes ambientes de produção, de modo a descartar a possibilidade de uma *generalização ilusória*.

5 Conclusão e Trabalhos Futuros

Este estudo demonstrou com sucesso o potencial das Redes Neurais Convolucionais para a predição de peso de tilápias a partir de imagens, alcançando um Erro Absoluto Médio (MAE) aceitável (45–50 g) e um Coeficiente de Determinação (R^2) superior a 0,90 na maioria dos casos. A análise detalhada dos erros mostrou que a maior parte das predições se mantém dentro de uma faixa estreita em torno dos valores reais, com desvio padrão dos erros absolutos em torno de 38 g e erro percentual médio inferior a 5%, o que é compatível com aplicações práticas em piscicultura intensiva.

Identificamos limitações importantes: (1) a ausência de dados físicos complementares, como o comprimento do peixe, que poderiam melhorar a precisão; (2) a variabilidade na qualidade das imagens (iluminação, posicionamento); e (3) o fracasso do *pipeline* de pré-processamento de imagens, que necessita de maior refinamento. Também foi observada uma leve tendência à superestimação dos pesos preditos, apontando para a necessidade de calibração adicional do modelo ou de técnicas de pós-processamento para correção de viés.

Como trabalhos futuros, sugerimos:

- Investigar métodos menos invasivos de obtenção de imagem (ex.: câmeras subaquáticas), eliminando a necessidade de remover o peixe da água e aproximando o sistema de um cenário de monitoramento contínuo.
- Integrar outras variáveis de entrada no modelo (ex.: comprimento e largura do peixe) para uma predição multimodal, combinando informações morfométricas explícitas com características extraídas automaticamente pelas CNNs.
- Explorar arquiteturas mais avançadas, como Redes Residuais (ResNets) e modelos leves (MobileNet, EfficientNet), bem como técnicas de *Transfer Learning*, de forma a avaliar o ganho de desempenho em comparação à arquitetura atual.
- Validar o modelo em bancos de dados maiores e mais diversificados, coletados em diferentes sistemas de produção e sob variadas condições de iluminação, para garantir a generalização e descartar a possibilidade de uma superestimação do desempenho.
- Investigar o uso de módulos de segmentação ou detecção de pontos anatômicos, capazes de isolar o contorno corporal da tilápia, reduzindo a influência de reflexos e ruídos de fundo sobre as predições.

Além dessas perspectivas, a implementação de técnicas de *Transfer Learning*, utilizando redes pré-treinadas em grandes bases de imagens naturais, surge como uma alternativa promissora para reduzir o tempo de treinamento e potencialmente melhorar a capacidade de generalização do modelo em cenários de poucos dados rotulados. Arquiteturas leves, como MobileNet ou EfficientNet, também se mostram interessantes para futura implantação em dispositivos embarcados, aproximando a solução proposta de aplicações em tempo real em unidades produtoras.

Do ponto de vista prático, recomenda-se, ainda, o avanço em diretrizes de captura de imagens em campo, como padronização de distância da câmera, controle básico de iluminação e manutenção de um fundo contrastante. Esses cuidados simples podem melhorar significativamente a qualidade das predições sem exigir modificações profundas na arquitetura da rede.

A automação desse processo tem potencial para melhorar significativamente o manejo na piscicultura, reduzindo custos, promovendo o bem-estar animal e aumentando a eficiência produtiva. Ao aproximar a tilapicultura de práticas de agricultura digital, este trabalho contribui com uma base experimental sólida sobre a qual soluções mais completas e integradas podem ser construídas nos próximos anos.

5.1 Limitações do estudo e perspectivas futuras

Embora os resultados obtidos sejam promissores, algumas limitações metodológicas devem ser consideradas. O objetivo principal deste estudo foi avaliar a viabilidade do uso de Redes Neurais Convolucionais para predição não invasiva do peso da tilápia a partir de imagens, priorizando a análise do comportamento do modelo sob diferentes configurações de treinamento e otimização. Dessa forma, o foco foi direcionado para a compreensão da estabilidade e da capacidade de generalização do modelo proposto, em vez da realização de uma comparação extensiva com múltiplas arquiteturas ou abordagens alternativas.

Adicionalmente, abordagens como *Transfer Learning*, arquiteturas leves (por exemplo, modelos da família *MobileNet*) e modelos estatísticos baselines não foram exploradas neste trabalho. A decisão foi tomada para manter o controle experimental sobre a arquitetura proposta e permitir uma análise clara do impacto de variáveis específicas, como otimizadores, pré-processamento e estratégias de aumento de dados. Ainda assim, reconhece-se que tais abordagens representam caminhos relevantes para futuras investigações, especialmente visando redução de custo computacional, melhoria da eficiência de inferência e ampliação da aplicabilidade prática em ambientes produtivos.

Por fim, destaca-se que o desempenho do modelo pode variar em função da distribuição de pesos e das condições de captura das imagens. Assim, estudos futuros deverão incluir análises estratificadas por faixas de peso, avaliações estatísticas mais aprofundadas, validação cruzada em cenários com maior disponibilidade computacional e a incorporação de informações adicionais, como segmentação mais robusta ou uso explícito de referências visuais (por exemplo, a régua presente nas imagens). Essas extensões poderão contribuir para aumentar a precisão do modelo e consolidar sua aplicação em sistemas reais de monitoramento na piscicultura digital.

Agradecimentos

Agradecemos ao núcleo PEIXEGEN pela disponibilização das imagens utilizadas neste estudo. Essas imagens foram essenciais para o desenvolvimento deste trabalho, permitindo treinar, validar e analisar os modelos de predição de peso das tilápias com rigor e confiabilidade. Reconhecemos a importância da parceria e do suporte fornecido, que contribuíram de forma significativa para a realização desta pesquisa.

References

[1] Associação Brasileira de Piscicultura (PeixeBR). O Brasil é o quarto maior produtor mundial de tilápia. <https://www.peixebr.com.br>, 2023. Acesso em: 20 jul. 2024.
[2] G.K. Iwama, A.D. Pickering, J.P. Sumpter, and C.B. Schreck. *Fish Stress and Health in Aquaculture*. Cambridge University Press, New York, 1997.

[3] Leonardo José Gil Barcellos, Sílvia Maria Guimarães de Sousa, and Viviane Mara Woehl. Estresse em peixes: fisiologia da resposta ao estresse, causas e consequências (revisão). *Boletim do Instituto de Pesca*, 26(1):99–111, 2000. URL <https://institutedepesca.org/index.php/bip/article/view/669/654>.
[4] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, Boston, 3 edition, 2016.
[5] Tom M. Mitchell. *Machine Learning*. McGraw-Hill Education, New York, 1997.
[6] Pedro Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55:78–87, 2012. doi: 10.1145/2347736.2347755. URL <https://dl.acm.org/doi/10.1145/2347736.2347755>.
[7] Weihong Ma, Xiangyu Qi, Yi Sun, Ronghua Gao, Luyu Ding, Rong Wang, Cheng Peng, Jun Zhang, Jianwei Wu, and Zhankang Xu. Computer vision-based measurement techniques for livestock body dimension and weight: a review. *Agriculture*, 14(2):306, 2024. doi: 10.3390/agriculture14020306. URL <https://www.mdpi.com/2077-0472/14/2/306>.
[8] Chang-Bok Lee, Han-Sung Lee, and Hyun-Chong Cho. Cattle weight estimation using fully and weakly supervised segmentation from 2d images. *Applied Sciences*, 13(5):2896, 2023. doi: 10.3390/app13052896. URL <https://www.mdpi.com/2076-3417/13/5/2896>.
[9] Alex Júnio da Silva Cardoso, Carlos Antonio Lopes de Oliveira, Eric Costa Campos, Ricardo Pereira Ribeiro, Gutierrez José de Freitas Assis, and Fabyano Fonseca e Silva. Estimation of genetic parameters for body areas in Nile tilapia measured by digital image analysis. *Journal of Animal Breeding and Genetics*, 138(6):731–738, 2021. doi: 10.1111/jbg.12551. URL <https://onlinelibrary.wiley.com/doi/10.1111/jbg.12551>.
[10] Jianlong Zhang, Yanrong Zhuang, Hengyi Ji, and Guanghui Teng. Pig weight and body size estimation using a multiple output regression convolutional neural network: A fast and fully automatic method. *Sensors*, 21(9), 2021. doi: 10.3390/s21093218. URL <https://doi.org/10.3390/s21093218>.
[11] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, Montréal, 2016.
[12] Sebastian Ruder. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016. URL <https://arxiv.org/abs/1609.04747>.